Comparing Mobility and Predictability of VoIP and WLAN Traces

Jeeyoung Kim, Ahmed Helmy
Department of CISE, University of Florida,
Gainesville, FL, USA
{jk2, helmy}@cise.ufl.edu

***Abstract*—How can we obtain *realistic* mobility models? This has been a question that many researchers have attempted to answer, mostly by analyzing existing WLAN traces. But in the future, will user on-line behavior change with the introduction of new mobile services and devices? We aim to investigate this issue in our study. In this paper, we analyze the mobility of a subset of users different than the general WLAN users; the VoIP users. These users are often mobile while on-line and their devices are *always 'on'*. We conjecture that their mobility is captured by the traces better than that of the WLAN users, and we expect them to be highly mobile users representing a trend for future mobile users. To that extent, we contrast the mobility of the VoIP, WLAN users, and three carefully selected sets of users across various metrics.**

We find that the fraction of time a VoIP user spends at a given AP is lower than WLAN users, indicating that the VoIP users are indeed more mobile than WLAN users. Also, the average and median number of access points visited for VoIP users are 4 to 8 times larger than that of the WLAN users. VoIP users cover a larger area range than WLAN users, indicating that VoIP users are physically more mobile. These findings point to significant difference in mobility characteristics between VoIP users and average WLAN users (commonly used for mobility modeling).

In order to examine whether this sharp contrast in mobility affects mobile networking protocols, we compare the performance of different classes of predictors across these different sets of traces. In particular, we evaluate the Markov O(1), O(2), O(3) and the LZ predictors. To our surprise, we find that the average prediction rate is over 60% for general WLAN traces while the prediction success rate drops below 25% for VoIP traces. Lessons learned in our study strongly suggest that both mobility modeling and location prediction should be re-visited in the context of future highly mobile users and devices.

Keywords-VoIP; WLAN; mobility; wireless traces; location prediction

I. INTRODUCTION

Realistic modeling of user mobility is one of the most critical research areas in wireless networks. Mobility data based on real human behaviors may give us the opportunity to improve wireless and mobile services for users in many ways. Currently, several mobility models are proposed based on the analysis of real WLAN traces [1,2,5,6,9]. However, the large collection of WLAN *usage* traces, seems to capture little *mobility* from the users. The average user is usually static while using the network, and exhibits a large *off time*.

In this paper, we focus on a subset of the wireless users, who use wireless VoIP devices. These users leave their devices *on* most of the time and the devices are light enough to carry and use while mobile. Hence, these users show a more mobile characteristic than laptop or other heavy device users while connected to the network. By analyzing these traces we aim to compare behavior of highly mobile VoIP users to the general WLAN users. This sheds light on the realism of WLAN trace-based models. We also aim to examine the effect of any differences on protocol performance, e.g., prediction protocols.

Particularly, we compare the mobility of VoIP user traces to whole WLAN traces (as used in previous studies) and also to some samples we have generated based on criteria that distinguish these samples as highly mobile compared to others. We use the metrics of prevalence, number of visited APs and acivity range defined in Section 3 to capture some of the main mobility characteristics of the users in our study. Our results clearly indicate that there is a significant difference between VoIP users and general mobile users, which strongly suggests revisiting mobility models of future *always-on* portable devices.

But does such dramatic contrast in mobility affect mobile networking protocols? In order to quantify such effect we examine the accuracy of several classes of mobility prediction protocols under various conditions of realistic mobility.

We compare these different sets of traces using several different predictors including the Markov O(1), O(2), O(3) and also the LZ predictor. Our experiments indicate that the Markov O(2) is the predictor with the highest accuracy among the four predictors and the LZ has the lowest. Surprisingly, all predictors perform quite poorly with VoIP users with an average of approximately 25% correct prediction rate, compared to 60% for the general WLAN users. These results prompt re-visiting of such algorithms for highly mobile users.

We provide guidelines and pointers to improve mobility modeling and prediction protocols for highly mobile users based on the lessons learned from this study. Based on such insight we plan to develop complete solutions to these problems in our future work.

The rest of the paper is organized as follows. In Section 2, we discuss related work and approaches.In Section 3, we outline our experimental setup along with background information on our data sets and metrics. In Section 4, we examine the difference of mobility between WLAN and VoIP users. In Section 5 we explore different predictors and the different prediction results between WLAN and VoIP users. Section 6 concludes the paper and discusses future work.

II. RELATED WORK

The related work lies in the areas of mobility modeling and mobility (and location) prediction. Among the numerous modeling techniques for mobility (random, synthetic, etc.) the most realistic is the trace-based mobility modeling. Many mobility modeling techniques so far were done based on analyzing the collective WLAN traces. Model T[5] and T++[6] are empirical registration models derived from the WLAN registration patterns of the mobile users. They are able to formulate the inter-dependence of space and time explicitly by a set of few equations.

In [1], Hsu et al. proposed a mobility model to capture time-variant user mobility. In this model, they define communities that are visited often by the nodes to capture the skewed location visiting preferences, and use time periods with different mobility parameters to create the periodical re-appearance of nodes at the same location. Hsu et al. [9] also looks into modeling generic WLAN users by identifying the mobility characteristics of individual users. In [7], Balazinska et al. studied user mobility patterns and introduced metrics to model user mobility from a four week trace collected in a large corporate environment. They also analyzed user distribution and load distribution across access points. Most of these works are directly based on WLAN traces which can be found under the MobiLib project [13] or the CRAWDAD project [14].

Interestingly, only a few researches were done on the VoIP trace. Kim et al. in [4] analyzed the VoIP trace from the Dartmouth WLAN trace and tried to come up with a mobility model by analyzing pause times, speeds, paths and locations of the users. But their paper does not analyze the difference of mobility and predictability between WLAN and VoIP traces. To the best of our knowledge there has not been any work done using predictors to compare the mobility of different users.

Song et al. in [3] investigated several domain-independent predictors for the location prediction on the WLAN trace, but did not define mobility characteristics or propose any techniques to construct the mobility model. Based on the comparison result, they gave some suggestions for the usage of the predictor on WLAN traces. There are a number of user mobility prediction algorithms [10, 11] in the current literature that target cellular networks. These predictors are used in a different setting and for different purposes (i.e. paging scheme [10], efficient handoff [10], resource reservation [11]). The characteristics and scale of the predictions mentioned in the above literature are different from what we are working on. The difference including, but not limited to, the fact that a cellular device showing up in a cell that is a long distance away is very low, thus it is bounded location-wise. Whereas, in our study the mobile user could easily log off and then log back on from a totally different location at a random time.

In our study, we use four predictors that have already been explored in existing literature [3] to verify the difference of the prediction accuracy due to mobility. The LZ predictor predicts in the case when the next symbol in the produced sequence is dependent on only its *current state*. The Order-k Markov predictor assumes that the location can be predicted from the current context which is the sequence of the k most recent symbols in the location history. The probability equation in the Markov Family considers how often the string of interest occurs in the entire input string.

III. EXPERIMENTAL SETUP

A. Data sets

We use the 3 year long Dartmouth movement trace [14] collected from 2001 to 2004 in our study. There are 7134 unique users and 623 different APs in this particular trace. While using this trace as our standard, general WLAN user base, we also extract our other data sets from this trace.

The VoIP data set we use in this work is a subset of the WLAN trace above and consists of 97 users. These are acquired by mapping the whole WLAN trace with a MAC to device type map, which is a list of all the MAC addresses mapped with the type of device it is by looking at the first three octets of the MAC addresses. Among these 97 users we observe 2 types of VoIP devices which are the Cisco7920 and Vocera devices. We have particularly chosen VoIP to measure the mobility of WLAN users since VoIP devices are always on the *on* state unlike other pocket PCs or PDAs that may easily go into hibernate mode or may even be turned on and off frequently.

Along with the VoIP data set we have generated test sets from the same trace in order to validate our findings. There are three test sets used in this work and they are all considered to be highly mobile users. Sample1 and Sample2 are both based on the number of APs visited. Sample1 is a collection of users who have visited 200 APs or more during the length of the trace and Sample2 is a collection of users who have visited more than 170 APs but less than 200. Sample3 is a collection of users who have covered the largest physical area during the length of the trace. This was done by studying the AP location file and calculating the area range that each user has covered. Each of these test set has approximately 100 users each. The following table 1 shows the different characteristics of the different data sets at a glance.

TABLE I. CHARACTERISTICS OF THE DIFFERENT DATA SETS EXTRACTED FROM DARTMOUTH MOVEMENT TRACE (2001-2004)

Labels	Number of users	Characteristic
WLAN	7134	All the users studied in this trace
VoIP	97	Users that use VoIP devices
Sample1	112	Users that have visited more than 200 APs
Sample2	98	Users that have visited more than 170 and less than 200 APs
Sample3	113	The users who has covered the largest physical area range

B. Metrics

How do we measure the difference of mobility that exists for different users? How do we say one user is more mobile than another? In order to answer these questions and quantify user mobility in order to compare and investigate users with different mobility levels, we come up with the following metrics: prevalence, number of visited APs and activity area.

Our evaluation metrics include prevalence, number of access points visited and the activity range. Prevalence of the traces indicates the *time that a user spends on a given AP, as a fraction of the total amount of time that they spend on the network*. The activity range comparisons is to analyze how wide an area each user has visited during each session, where the definition of session is *the time duration from the user connecting to the network to when they disconnect or disappear from the network*. The activity range is defined as *the smallest square area which can cover all the access points the user visited in an activity*. We also analyze the activity range by each day (24 hours) to see what the activity range is in different times. To compare the performance of different predictors we use the prediction accuracy metrics which define the percentage of correct prediction for each user in the following section.

IV. MOBILITY COMPARISON

In our work, we compared the mobility characteristics of WLAN traces and VoIP traces from several different aspects. The evaluation metrics include prevalence, the number of access points visited by the users and the activity range where a user has been active in. The results of the comparison for each of our evaluation metrics is listed and shown as follows.

A. *Prevalence*

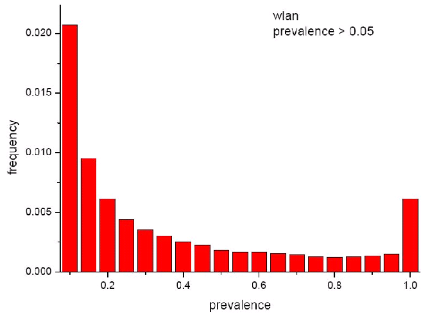

Figure 1. Prevalence of WLAN trace

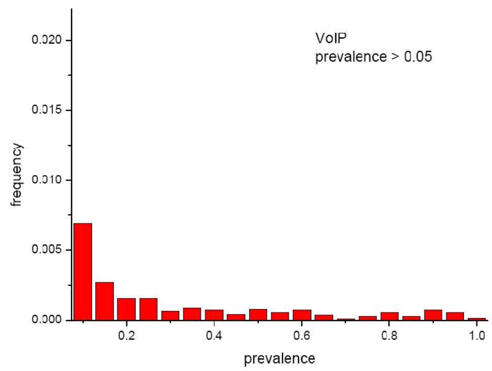

Figure 2. Prevalence of the VoIP trace

Prevalence is one of the mobility metrics proposed in [7], which indicates *the time that a user spends at a given AP, as a fraction of the total amount of the time that they spend on the network*. Higher prevalence means user spent more time on such an AP, and thus less mobile. Figure 1 and 2 show that VoIP users are more mobile than WLAN users, since the bar is lower. Especially for the most right bar which indicates prevalence higher than 0.95, the WLAN is much higher than VoIP. This means there are larger portion of users in WLAN who spent most of their time on only one AP than that in VoIP.

B. *Number of access points visited*

Figure 3 and 4 shows the number of access points visited distribution CDF by WLAN and VoIP users. This clearly shows VoIP users visited much more access points than WLAN users. The average number of access points that the VoIP users visited is about 4.1 times than that of the WLAN users while the median number of access points that the VoIP users visited is about 7.7 times than that of WLAN.

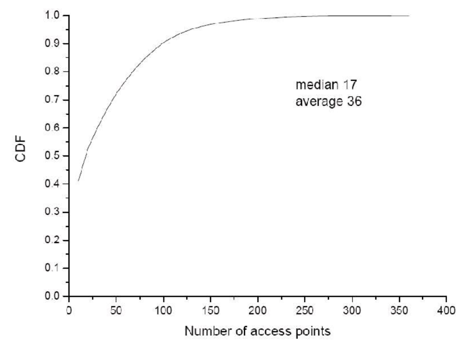

Figure 3. Number of Access Points visited in the WLAN trace

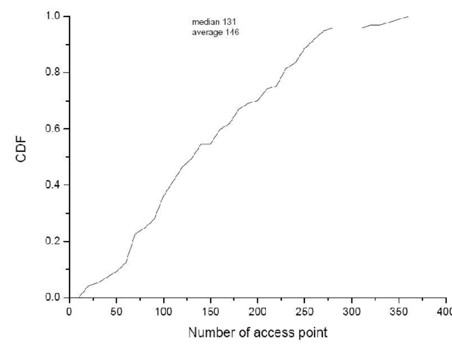

Figure 4. Number of Access Points visited in the VoIP trace

C. *Activity Range*

Activity range is defined as *the smallest square area which can cover all the access points the user has visited in an activity* as shown in figure 5.

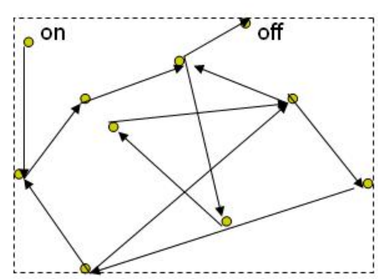

Figure 5. Activity Range

Figure 6 and 7 shows the activity range distribution for WLAN and VoIP users. The percentage of VoIP users having a larger area of activity range is higher than that of the WLAN users who, most of the time tends to stay in a very limited area.

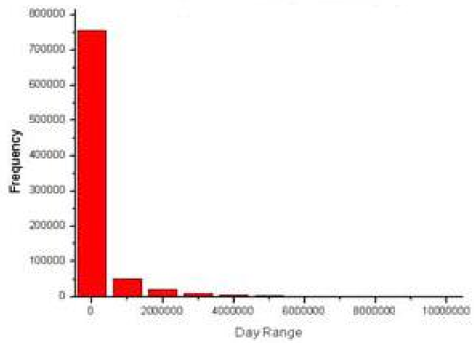

Figure 6. Activity Range Distribution of WLAN

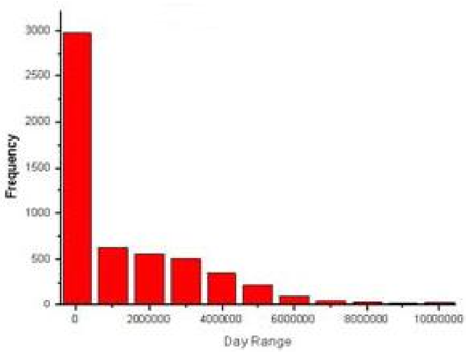

Figure 7. Activity Range Distribution of VoIP

V. PREDICTABILITY COMPARISON

To study the effect of the sharp contrast in mobility and behavioral characteristics between VoIP and other WLAN users on networking protocols we analyze a set of well known prediction algorithms with the various sets of traces we have in our study.

We have run the Markov O(1), O(2) and O(3) predictors along with the LZ [3] predictor for each of the test sets we have, and also for the VoIP trace set and the whole body of the WLAN trace. We also compared the accuracy of all four predictors with the VoIP trace data to see which one has the best performance. Accuracy is measured as percentage of correct predictions of the next AP to visit. As shown in figures 8 through 12, the ***WLAN trace always had the best prediction accuracy for all the predictors with an average of about 60% accuracy. The VoIP trace, by contrast, had the worst prediction accuracy for all of the predictors with an average of approximately 25% accuracy.*** From these graphs we see that the best accuracy can be no more than 80% for VoIP users, while more than 95% accuracy for WLAN users.

When we were first conducting our experiment, we expected that the range of the physical area that each user covered would be a better criteria to measure mobility than the number of APs visited since we consider a person to be more mobile when that person covers more ground. Hence, we expected that sample 3 would return a very bad prediction accuracy. Surprisingly, sample 3 always exhibits performance between of the other two samples (1 & 2), which indicates that the users that covered larger areas physically most likely have visited an average of 200 APs during their lifetime.

To explain this result, intuitively the users that had visited less APs also had a better prediction rate than that of the users who had visited more APs. The difference of the prediction accuracy between the two samples are always around 10% near the median.

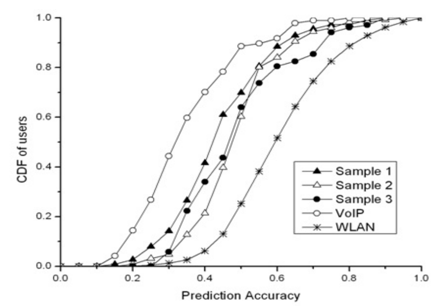

Figure 8. Accuracy of Markov O (1) Predictor

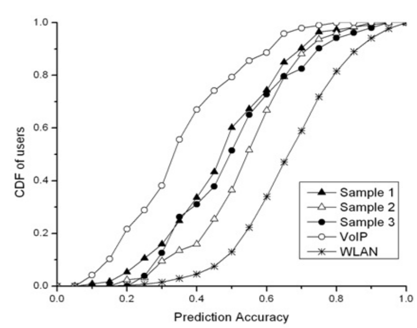

Figure 9. Accuracy of Markov O(2) Predictor

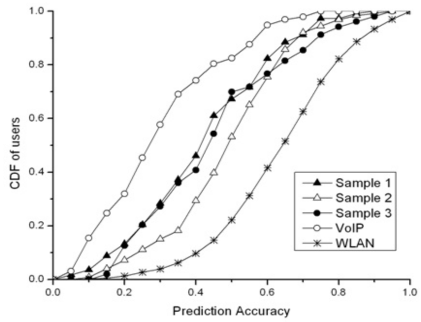

Figure 10. Accuracy of Markov O(3) Predictor

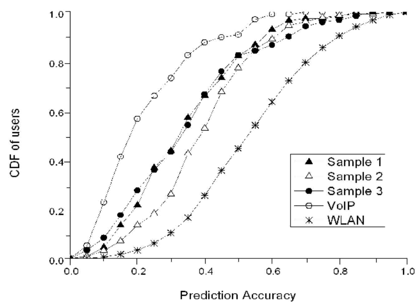

Figure 11. Accuracy of LZ Predictor

As for the comparison of the predictors on the VoIP data set, the LZ predictor showed the worst prediction rate and the Markov O(2) showed the best prediction accuracy by a very minimal difference from the Markov O(1). Markov O(3) did not show a good prediction and these results indicate that a larger data structure and higher complexity does not help in making better predictions. However, the four predictors that are used in this work do *not* provide good prediction for the VoIP data set.

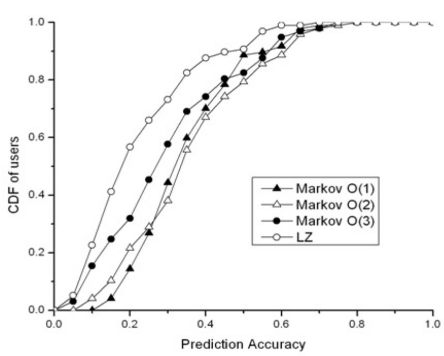

Figure 12. Comparison of Predictability on VoIP Trace

VI. CONCLUSION AND FUTURE WORK

Our findings open the door for revisiting mobility modeling and improved prediction of highly mobile users. We can see from our findings that whatever protocols and services (i.e. prediction) that were developed for the normal WLAN user can change dramatically in the environment of highly mobile users. We plan to design a better predictor for "highly mobile" users, especially the VoIP traces. Our plan includes investigating domain-specific knowledge, regressions, schedules and repetitive or preferential user behavior. We shall also examine the adequacy of WLAN trace based mobility models for highly mobile and VoIP users, that are likely to increase in the future.

ACKNOWLEDGMENT

We thank Yi Du and Mingsong Chen who have helped in establishing the fundamental findings in this paper.

REFERENCES

[1] W. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, *"Modeling Time-variant User Mobility in Wireless Mobile Networks"*, Proceedings of IEEE Conference on Computer Communications, May 2007, Anchorage, Alaska

[2] C. Tuduce, T. Gross, *"A mobility model based on WLAN traces and its validation"*, Proceedings of INFOCOM 2005: Miami, FL, USA 664-674

[3] L. Song, D. Kotz, R. Jain and X. He, *"Evaluating location predictors with extensive Wi-Fi mobility data"*, Proceedings of IEEE INFOCOM 2004, Hong Kong, China

[4] M. Kim, D. Kotz, and S. Kim, *"Extracting a mobility model from real user traces"*, Proceedings of IEEE INFOCOM 2006, Barcelona, Spain

[5] R. Jain , D. Lelescu , M. Balakrishnan, *"Model T: an empirical model for user registration patterns in a campus wireless LAN"*, Proceedings of the 11th annual international conference on Mobile computing and networking, August 28-September 02, 2005, Cologne, Germany

[6] D. Lelescu , U. Kozat , R. Jain , M. Balakrishnan, *"Model T++:: an empirical joint space-time registration model"*, Proceedings of the seventh ACM international symposium on Mobile ad hoc networking and computing, May 22-25, 2006, Florence, Italy

[7] M. Balazinska and P. Castro, "*Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network*" in International Conference on Mobile Systems, Applications, and Services, May 2003

[8] T. Henderson, D. Kotz, I. Abyzov, *"The changing usage of a maturecampus-wide wireless network"*, Proceedings of the MOBICOM 2004: 187-201

[9] W. Hsu and A. Helmy, *"On Modeling User Associations in Wireless LAN Traces on University Campuses"* The Second International Workshop on Wireless Network Measurement (WiNMee 2006), Boston MA, Apr. 2006.

[10] H. Zang and J. C. Bolot, *"Mining Call and Mobility Data to Improve Paging Efficiency in Cellular Networks"*, Proceedings of the MOBICOM 2007: 123-134

[11] J. Chan and A. Seneviratne, *"A Practical User Mobility Prediction Algorithm for Supporting Adaptive QoS in*

Wireless Networks", Proceedings of IEEE International Conference on Networks, Sep. 1999, Brisbane, Australia

[12] F. Chinchilla, M. Lindsey, M. Papadopouli, "Analysis of wireless information locality and association patterns in a campus", Proceedings of the IEEE INFOCOM 2004, Hongkong, China

[13] http://nile.cise.ufl.edu/MobiLib/

[14] http://crawdad.cs.dartmouth.edu/